

Entailment as Plural Modal Anaphora

The phenomenon. Entailment is a phenomenon sensitive to the *content* of expressions as opposed to other components of their meaning, e.g. the Partee marble examples differ in *meaning* (context-change potential), but they are content equivalent. Other content-driven phenomena (often involving entailment relations) are answerhood conditions, propositional attitude reports, presupposition resolution (e.g. clefts and factives) and rhetorical relations (e.g. clarification requests and check moves in dialogue)

At the same time, it is well known that the *content* of natural language expressions is dependent on context and meaning: we need to access discourse referents set up in a particular context to determine the content of pronouns and other anaphors, e.g. (1) below entails (2) when uttered on a Thursday in a discussion about John, but not otherwise.

(1) **He** came back **three days ago**. \models (2) **John** came back **on a Monday**.

Moreover, we need to take into account the anaphoric dependencies between premises and conclusion to be able to determine their contents, as the pairs (3)-(4), (5)-(6) and (7)-(8) below show. The pair (7)-(8) is interesting because we need an account of *modal anaphora* between (8) (i.e. *would*) and (7) (i.e. *might* and *would*) to ascertain that entailment between them obtains.

(3) **A man** came in. \models (4) **He** entered. (5) **Every man** saw a woman. \models (6) **They** noticed them.
 (7) ¹ A wolf **might** come in. ² It **would** see John. \models (8) It **would** notice him.

In sum, to be able to determine entailment relations (or answer questions, resolve presuppositions etc.), we need to *access the contents* of natural language expressions *within* discourse / dialogue, while at the same time account for the fact that: (a) determining content is context sensitive and (b) premises supply the context for the conclusion.

The proposal. The central observation underlying our proposal is that modal anaphora (see the modal subordination example in (7) above) provides a paradigm for accessing contents and determining relations between them. Building on previous work on modal and plural individual anaphora ([4] and [5]), we describe a new dynamic system in which we can represent both the *static contents* and the *dynamic meanings* of utterances in discourse.

Following [2], the dynamic system is formulated in many-sorted type logic; the basic types are t (truth-values), e (individuals), s (modeling variable assignments) and w (possible worlds). For simplicity, we ignore the temporal and eventuality domains. Just as in [5], information states I, J etc. are sets of assignments (i.e. type st). An individual discourse referent (dref) u (type se) stores a plurality with respect to an info state I , abbreviated $ul := \{x_e : \exists i_s \in I_{st} (x=ui)\}$, where the subscripts on terms indicate their types. A modal dref p (type sw) stores a proposition (a set of worlds), $pl := \{w_w : \exists i_s \in I_{st} (w=pi)\}$.

Encoding pluralities and propositions in this way and not via drefs for sets (their type would be $s(et)$ and $s(wt)$ respectively) allows us to capture *structured* intersentential plural anaphora: e.g. (5) above entails (6) only if, for each man m that saw a woman n , we assert that m noticed n . Entailment does not obtain if man m_1 saw woman n_1 and m_2 saw n_2 , but we interpret (6) as asserting that m_1 noticed n_2 and m_2 noticed n_1 . The plural info states (type st) allow us to store and pass on this *distributive* structure: for each $i_s \in I_{st}$, we require that the man in i saw (hence noticed) the woman in i .

Sentence *contents* (propositions) are represented as modal drefs. Sentence *meanings* are Discourse Representation Structures (DRSs), i.e. relations between info states (type $st(stt)$). For example, the content of (2) above is the proposition $\{w : come_w(john)\}$, i.e. the *maximal* set of worlds w in which John comes back. We add a *max* operator over modal drefs (compare the maximality in the individual domain argued for in [5]) to be able to extract and store this proposition. Thus, the meaning of (2) is given in the DRS in (9) below and the dref p encodes the content of (2); we define the lexical relation and the proper name dref in (10) and *max* in (11) below. Max^p can be thought of as dynamic λ -abstraction over possible

worlds: it introduces a new modal dref p (symbolized as $I[p]H$) and makes sure that each and every world in pJ satisfies DRS D ; the set pJ is maximal because any other set of worlds satisfying D , stored as pK , has to be included in pJ .

(9) \mathbf{max}^p ($[\mid \text{come}_p\{\text{John}\}]$) (10) $\text{come}_p\{\text{John}\} := \lambda I. \forall i \in I(\text{come}_{pi}(\text{John}_i)); \text{John} := \lambda i. \text{john}_e$.

(11) \mathbf{max}^p (D): $= \lambda IJ. \exists H (I[p]H \ \& \ DHJ) \ \& \ \forall K (\exists H (I[p]H \ \& \ DHK) \rightarrow pK \subseteq pJ)$

All this machinery is independently needed to properly represent modal anaphora in (7): discourse (7) is true if there is an epistemic possibility p in which a wolf u comes in and sees John and, *in addition*, any epistemic possibility p in which a wolf u comes in is such that the wolf u sees John (see [3]). So, to properly represent (7) – see (12) below, we need *maximality* (we consider *any* epistemically accessible world in which a wolf comes in) and also *distributivity*: just as in (5)-(6), we establish a *structured* correspondence between each world w and the wolf x that enters in w (and which can vary from world to world). If wolf x_1 enters in world w_1 and x_2 in w_2 , sentence (7²) requires w_1 to be such that x_1 (and not x_2 !) sees John.

(12) \mathbf{max}^p ($[\text{ul wolf}_p\{u\}, \text{come_in}_p\{u\}, p \subseteq p_0]$; $[\mid \text{see}_p\{u, \text{John}\}]$).

The modal verb *might* introduces and maximizes p ; the dref p_0 encodes the contextually specified set of epistemically accessible worlds. The modal anaphor *would* refers back to and elaborates on p . These ingredients are sufficient for the definition of entailment in (13) below.

(13) A DRS D interpreted relative to modal dref p *entails* a DRS D' interpreted relative to the same dref p with respect to an input info state I *iff* for any state K such that \mathbf{max}^p (D) IK , there is an output state J such that \mathbf{max}^p (D') KJ and $pK \subseteq pJ$.

We use two info states K and J to store the *different* contents of the premise and of the conclusion, i.e. pK is the content of D with respect to context I and pJ is the content of D' with respect to context K . We use the same modal dref p because the premise and the conclusion are part of the *same* argument. If we had used two different drefs to store the contents of D and D' , we would have incorrectly predicted that entailment obtains between *John might come in. It would see Mary* and *He (John) notices her (Mary)*, where the premise is interpreted relative to a hypothetical scenario (modal dref) of John seeing Mary, while the conclusion (in indicative mood) is interpreted relative to the actual scenario / world.

This definition recaptures in a dynamic setup the intuitive appeal of 'entailment is content / proposition inclusion', as modeled in possible-worlds semantics. Also, it accounts for context sensitivity (the input state I) and for structured anaphora between premises and conclusion (the intermediate state K). The premise-conclusion anaphoric connections are also captured by the definition of entailment in [1]: $\forall I \forall K (DIK \rightarrow \exists J (D'KJ)$). However, besides effacing the meaning-content distinction, this definition is too strong: the universal quantification over the input state I rules out the entailment between (1) and (2) above. The version without quantification, i.e. $\forall K (DIK \rightarrow \exists J (D'KJ)$), is too weak: if sentences *John is happy* and *John is tired* are interpreted in a context in which John is both happy and tired, we incorrectly predict that entailment obtains. The definition in (13) is intermediate in strength: we preserve the input context I with the exception of the value assigned to the modal dref p .

The definition of *truth* underlying the entailment definition in (13) above is: a DRS D interpreted relative to a modal dref p is *true* with respect to an info state I and a world w *iff* there is an output state J such that \mathbf{max}^p (D) IJ and $w \in pJ$.

References: [1] Groenendijk, J. & M. Stokhof 1991. Dynamic Predicate Logic, *Linguistics and Philosophy* 14, 39-100; [2] Muskens, R. 1995. Tense and the Logic of Change, *Lexical Knowledge in the Organization of Language*, U. Egli et al (eds.), Benjamins, 147-183; [3] Roberts, C. 1989. Modal Subordination and Pronominal Anaphora in Discourse, *Linguistics and Philosophy* 12, 683-721; [4] Stone, M. 1999. *Reference to Possible Worlds*, RuCCS Report 49, Rutgers University; [5] van den Berg, M. 1996. *Some aspects of the Internal Structure of Discourse. The Dynamics of Nominal Anaphora*, PhD Dissertation, University of Amsterdam.